DOCUMENT RESUME

ED 383 768 TM 023 316

AUTHOR Rossman, Gretchen B.; Salzman, Julianne
TITLE Evaluating Inclusive Education Programs: A Survey of Current Practice. Preliminary Analysis and a Typology.
SPONS AGENCY City Univ. of New York, NY. National Center on Educational Restructuring and Inclusion.
PUB DATE Apr 95
NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classification; Databases; *Educational Practices; Elementary Secondary Education; *Evaluation Methods; Heterogeneous Grouping; *Inclusive Schools; Literature Reviews; Mail Surveys; Mainstreaming; Outcomes of Education; Parent Participation; *Program Evaluation; Research Methodology; *Special Education; Telephone Surveys

ABSTRACT

Initial steps in a long-term effort to identify and analyze evaluations of inclusive education programs are discussed. Three activities have been initiated to survey current evaluation practice. A literature search revealed that "inclusion" is not yet a description for the ERIC system, but that some papers have been published on the topic. A telephone survey of state directors of special education has begun, with 10 interviewed to date. A mail survey has begun of schools and districts identified as part of the National Center on Educational Restructuring and Inclusion database on inclusive programs. When the three efforts are completed, a report will be written to suggest principles for comprehensive evaluation of inclusive education programs. Programs can be classified by purpose, complexity, scope, population served, and duration. A variety of evaluation designs and methods are being employed. Most evaluations studied so far have focused on student outcomes, specifically academic and social gains. Support from parents, staff, and students is another focus of many evaluations. (Contains 10 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 383 768

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
GRETCHEN B. ROSSMAN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Evaluating Inclusive Education Programs: A Survey of Current Practice

Preliminary Analysis and a Typology

Gretchen B. Rossman
Julianne Salzman

University of Massachusetts at Amherst

Paper presented at the annual meetings of the American Educational Research Association, San Francisco, April, 1995. This paper was commissioned by the National Center on Educational Restructuring and Inclusion; no official endorsement should be inferred.

TM023316

BEST COPY AVAILABLE

Evaluating Inclusive Education Programs:
A Survey of Current Practice

Preliminary Analysis and a Typology[1]

Gretchen B. Rossman
Julianne Salzman
University of Massachusetts at Amherst

April, 1995

A call heard from federal officials, state directors of general and special education, and local administrators is that there is insufficient systematic evaluation of programs designed to more fully include children with disabilities in the general classroom. While there have been discrete assessments of, for example, teacher inservice needs, student social and academic progress, and parental attutides, a search of the ERIC databases reveals few comprehensive, wholistic evaluations of inclusive education programs. It is entirely likely that many such evaluations are conducted and not published or disseminated through standard media; this certainly limits their accessibility to other evaluators or program planners. This paper discusses the initial steps in a long-term effort to identify and analyze evaluations of inclusive education programs.

Three activities have been initiated to survey current evaluation practices. To date, we have made an initial search of the ERIC databases looking for program evaluations of mainstreaming, integration, and inclusion. As as aside, it was interesting to discover that the term *inclusion* is not yet a legitimate ERIC descriptor. Thus far, this search has yielded 17 usable documents dating from 1990. Second, we have begun a telephone survey of state directors of special education in all 50 states, Puerto Rico, and the Virgin Islands. To date, we have spoken with 10 directors whose responses to our questions ranged from outrage that we would use the term *inclusion* (!) to strong help and interest in this effort. The third activity will be a mail survey to all schools and districts identified as part of the National Center

---

[1] An earlier version of this paper was presented at an invitational conference at Wingspread, Racine, WI, in April, 1994.

on Educational Restructuring and Inclusion database on inclusive programs. This will be conducted over the summer and fall of 1995 and incorporated into a revised version of this paper.

This paper is divided into two major sections. First, we present a classifying scheme for organizing the program evaluations identified thus far, as well as those to be identified as this search continues. This is broken down into features of the inclusive programs themselves, and features of the program evaluations. We provide examples from data collected to help flesh out this scheme. The second section describes findings of the evaluations, highlighting important or unusual concepts or conditions that support inclusive programs. When the three survey efforts are completed (probably by the end of the summer), a final section will be written that suggests principles for comprehensive evaluations of inclusive education programs.

The Classification Strategy

Two fundamentally different strategies can be used to categorize the evaluations identified to date, as well as others that we identify. First, they can be organized by *program features* which focuses on program purpose, scope, target population, duration, and complexity. Second, the evaluations can be categorized according to *features of the evaluation*. Here the analysis focuses on design complexity, evaluation methods utilized, and role groups from whom data were gathered (or unit of analysis. Each strategy highlights certain aspects of the program and its evaluation: the first stresses program description and findings from the evaluation of that program; the second, more methodological, emphasizes evaluation methods *per se* and permits syntheses about methods across a number of programs. Each has merit.

Our intent is to build a database that will articulate with the National Center on Educational Restructuring and Inclusion database which allows us to flip from one categorizing scheme to the other. That is, if we wanted to retrieve information on evaluation results for school-based programs for children with so-called severe disabilities, we could easily do so by programming in those descriptors. Similarly, if we wanted to identify those

program evaluations relying on surveys of peers in inclusive classrooms so that those instruments could be shared with interested parties, we could also do this search relatively easily. For the preliminary analyses presented here, we describe each feature in turn, providing examples from the identified evaluations.

Program Features

Purpose. Inclusive education programs have many different purposes, some quite singular, others more multi-faceted. For example, a program at one high school in Texas is intended to more fully serve 120 students described as having learning disabilities in general education classrooms (Chase & Pope, 1993); this purpose is relatively singular and straight-forward. In contrast are the statewide systems change grants that identify multiple purposes for the grants to achieve.

Complexity. Related to purpose, complexity captures the relative simplicity or complexity of the program. For example, a program serving students labelled as having challenging behavior is designed with collaborative consultation between general education teachers and specialists as its sole (or most notable) feature (Burrello & Wright, 1993); this program would be categorized as conceptually less complex than one that incorporates consultation, building-based planning teams, leadership training, parental support, peer coaching, and cooperative learning models (Rogan & Davern, 1992). More complex programs should provide complex results that describe particular successes and "worries" that would be instructive to others.

Scope. Programs of inclusive education can be sorted by scope. By this we mean whether the project was designed for a single school, a set of schools (perhaps at different levels in the system), an entire district (or, for example, all middle schools within a district), a set of districts within a state, or all districts within a state. This feature captures evolution in scope over time, also. The ERIC and other library searches have identified several inclusive programs serving students in one school (e.g., Chase & Pope, 1993; Burrello & Wright, 1993; Co-teaching, 1991); others designed to provide

more inclusive educational experiences for students in several schools within a district (Rogan & Davern, 1992; Marwell, 1990); those serving several districts (identified through various processes) across a state (Christmas, 1992; McDonnel et al., 1991; Ferguson et al., 1992); as well as the statewide systems change grants administered at the state level for several demonstration districts within the state.

Population served. A fourth programmatic feature allows us to distinguish the specific population served, usually by the disability with which students are labelled. This feature may be singular or blur the distinctions among children; that is, one program may focus on students labelled as having learning disabilities (e.g., Chase & Pope, 1993), while another might target all students previously served in substantially separate classrooms (we have not yet identified a program like this). Other programs may focus on teachers or paraprofessionals (e.g., Christmas, 1992). This feature helps distinguish between programs in useful ways, providing information on the successes and challenges of programs for specific populations of children or those who serve children.

Duration. A fifth feature of interest is the duration of the program. It would be useful to know, for example, that a district had implemented a program serving students with so-called severe disabilities fifteen years ago. That this program is operational and successful, and has met challenges and evolved over time, would yield different information than from a program that was at its inception.

Evaluation Methodology Features

This strategy focuses on the program evaluation itself, seeking to describe its salient features so that others can learn how a comprehensive (or not so comprehensive!) evaluation was conceived and conducted. By classifying the evaluations according to design, methods, instrumentation, and sample or unit of analysis, this strategy permits inspection of the evaluation itself. Furthermore, through this strategy, rich and complex designs can be identified and described; particularly insightful or creative methods can become

accessible; and useful instruments can be identified and categorized. When sorted with the program classifications above, this typology can generate examples of evaluations that are interesting and sound methodologically that have focused on specific programs of interest.

Design. Evaluation designs range from simple summative, "one-shot" evaluations that rely solely on a survey of one role group to complex formative and summative designs that use multi-method approaches, gathering data from a number of participant groups through a variety of methods. An example of a simple design is the evaluation conducted of an inclusive program for students labelled as behaviorally challenging (Burrello & Wright, 1993). Although the published report contains incomplete information about the evaluation, the data presented were derived from a survey of all staff regarding their perceptions about the success of this program. A second example of a simple design is found in the evaluation of a co-teaching program where a survey with both forced choice and open-ended items was administered to a number of participants and stakeholders (Co-teaching, 1991). A more complex design was used in the evaluation of the Syracuse City program to more fully include students described as having severe disabilities (Rogan & Davern, 1992). This design featured both process and formative evaluation components, and gathered data from a variety of role groups.

Methods. Flowing directly from the complexity of the design are the methods used in the evaluations. As noted above, evaluations may rely on one method -- for example, a survey -- to assess the effectiveness or success of the inclusive program. Others rely on multiple methods to triangulate among data sources. One example of the use of multiple methods is the evaluation of the Madison, Wisconsin, program to integrate students described as being mentally retarded[2] (Marwell, 1990). The methods used were interviews with a number of role groups, questionnaires, and sociometric analyses of classrooms. A second example of multiple methods is the published evaluation of the Utah

---

[2] This term is used in the original work.

elementary integration model focusing on students described as having severe disabilities (McDonnell et al., 1991). This evaluation relied on measures of program implementation, students' adaptive behavior through a validated instrument, level of integration by time analyses, and a survey of participating teachers.

Instrumentation. Of the evaluations identified to date, about half relied on formal instruments to assess their programs. Some of these are included in the reports; others can be extrapolated from the evaluation findings presented; and yet others are not retrievable through the reports. Our plan is to identify those instruments used, gather them into some sort of compendium of evaluation instruments, and make them available (with proper citation and permission from the designers) to people who are interested. To date, we have identified questionnaires of various role groups (e.g., teachers, parents, students, paraprofessionals) regarding their perceptions about the inclusive program; highly structured observation protocols for use in inclusive classrooms; and interview guides or protocols for use with a variety of role groups.

Sample. The samples from which evaluation data are gathered vary enormously in the work identified thus far. Some of this is a function of the complexity of the program itself: the more complex programs seek evaluative data from a number of samples of people affected by the program. An example comes from the evaluation of an inclusion initiative run out of the University of Oregon (Ferguson, 1992) that sampled students participating in the program (both disabled and nondisabled), teachers, classrooms, and schools. Programs more limited in complexity tend to sample only one role group, as in the evaluation of a program for the inclusion of students described as behaviorally challenged (Burrello & Wright, 1993) that sought evaluative data from staff as the only sample.

As we build the evaluation database, the above features of both inclusive programs and their evaluations will be used to code and sort the evaluations. Of further interest, however, are the findings of these

evaluations which we discuss next.

Findings

The findings reported in the evaluations identified thus far vary according to the questions each evaluation pursued. Some focused on student outcomes, others on staff perceptions of the inclusive program, and yet others on levels of implementation; some, of course, posed a set of questions covering a variety of potential processes and outcomes. The findings are clustered into the following six categories: student outcomes, parent support, student support, staff support, implementation, and overall effects.

Student Outcomes

Most evaluations identified thus far include questions on how students fare in more inclusive programs. Some asked discrete questions about social and academic learnings; others focused on time spent in inclusive classrooms; and yet others analyzed social gains.

Time. The evaluations analyzed thus far have found that students labelled as having learning disabilities spent more time in the general classroom than previously, as a result of the inclusive program; this was most dramatic for those coming from substantially separate classrooms (Chase & Pope, 1993). In the Utah program serving students labelled as having severe disabilities, after implementation of an inclusive program, time spent with nondisabled peers rose (McDonnell et al., 1991).

Academic gains. The evaluations generally suggest that students in inclusive programs made academic gains regardless of labelled disability. For example, students described as having learning disabilities made academic gains as reflected in gains on criterion-referenced testing and on report cards (Chase & Pope, 1993). Integrated students described as having severe disabilities, moreover, had greater success in achieving 8 IEP goals than did matched students in traditional programs (Ferguson, 1992). A co-teaching program intended to support students labelled with mild disabilities fostered much growth among mainstreamed students, particularly in terms of social skills and attitudes towards education (Co-teaching, 1991). This evaluation

concluded, moreover, that the project did not appear to have slowed down or curtailed educational process available to regular students (Co-teaching, 1991).

Social gains. The evaluations also found positive changes in social learnings for students in inclusive programs; some focused on the included students only, others on their nondisabled peers. The evaluation of a program for students described as behaviorally challenging found significant changes in the self-esteem (Burrello & Wright, 1993). Similarly, integrated students described as mentally retarded were generally accepted by classmates, with 61% receiving sociometric ratings near the mean and 29% in the socially "neglected or rejected" range (Marwell, 1990). This same evaluation found that general education teachers identified positive social effects for nondisabled students, as well (Marwell, 1990). In the evaluation of a statewide program focused on expanding the role of nonmandated aides to support students with disabilities in the general classroom found that the target students seemed integrated and accepted (Christmas, 1992). In the Utah program for the inclusion of students labelled as having severe disabilities, students demonstrated statistically significant gains ($p<.001$) on all subparts of a comprehensive social skills assessment (McDonnell et al., 1991). In one of the more interesting comprehensive evaluations of a program designed to include students described as having severe disabilities in the general classroom, the evaluators found evidence of repeated instances of *"bubble kids"* -- kids in the regular classroom who were integrated but isolated or separated (Ferguson, 1992).

Parent Support

Several evaluations focused on parent support for the inclusive program or parent attitudes towards inclusion generally. Parent support was described as overwhelmingly enthusiastic for the inclusive program for students described as having learning disabilities (Chase & Pope, 1993). Similarly, parents of integrated students labelled as mentally retarded were generally satisfied with the inclusive program, with 85% saying they would choose an

integrated program over a more traditional model (Marwell, 1990). This same evaluation found that 90% of the parents of students with disabilities believed that academic and behavioral standards had been maintained in the inclusive program (Marwell, 1990).

A co-teaching program to support students described as having mild disabilities received strong support from parents surveyed (Co-teaching, 1991). In elaborating on this finding, the evaluation indicated that parents were overwhelmingly supportive of the project, would like it expanded, and felt it had a positive impact on children in terms of attitudes towards self, peers, and school (Co-teaching, 1991).

Student Support

Many of the evaluations focused on student perceptions about the inclusive program, some targeting the students being included, others their nondisabled peers. The evaluation of a program serving students described as having learrning disabilities found that student support high, but fails to mention whether this was all students, targeted students, or some combination (Chase & Pope, 1993). The co-teaching program's evaluation also found strong support among students; in this instance, the authors describe that this sample includes both students with disabilities and non-disabled students in co-teaching classrooms who were surveyed (Co-teaching, 1991).

Staff Support

Many evaluations seem to find the simple survey of participating teachers an easy way to generate some evaluation data. While this is not ideal (as will be outlined in the third section of the revised paper), it does provide a perspective on the inclusive program. Analysis to date suggests that teacher support varies somewhat, but most responses are quite positive and supportive of inclusive programs. In a program serving students with learning challenges, teacher support ranged from excellent to fair (Chase & Pope, 1993). In a staff development program to provide training and support in collaborative consultation for students described as behaviorally challenging, 89% of staff rated their training in collaborative consultation as above

average to outstanding; 89% rated their involvement in collaboration meetings as above average to outstanding; 90% rated the collaborative teams as above average to outstanding; and 77% rated the developing joint ownership of student problems as reducing teacher anxiety as above average to outstanding (Burrello & Wright, 1993).

Similarly, in a staff development effort targeting district and building leadership to sensitize and build support for an inclusive program for students described as having severe disabilities, participants rated the leadership institutes very high (Rogan & Davern, 1992). Moreover, the co-teaching program to support students described as having mild disabilities received strong support from the teachers surveyed (Co-teaching, 1991). This program, moreover, concluded that the implementing teachers were very enthusiastic, felt they had gained professionally, and would prefer to continue in co-teaching classrooms. Informal statements from teachers indicated they had grown in teaching skills and in appreciation of their team partner's educational role, while non-project teachers were aware of project, unanimously in favor of further integration, and receptive to teaming (Co-teaching, 1991).

In the Utah program serving students labelled as having severe disabilities, the general education teachers were generally satisfied with the program although they disagreed on whether student with severe disabilities required a lot of extra attention from the homeroom teacher (McDonnell et al., 1991). An unusual finding emerged from the evaluation of a comprehensive program designed to support and foster the full inclusion of students described as having severe disabilities. This was that teachers demonstrated what the evaluator describes as "professional preciousness", that is, a tendency to define problems in ways that demand the available resources rather than more creatively or divergently (Ferguson, 1992).

Focusing on a different role group, the evaluation of the program to expand the role of nonmandated aides to support students with disabilities in the general classroom where the survey of aides found that they believed that

the inclusion project was worthwhile, they would participate in it again, that it was a benefit to all, but that they would have liked more training and visits to other inclusion projects. They also identified a need for more planning time (Christmas, 1992). This same evaluation also found that general education and special education teachers' opinions paralled those of the aides with the exception of the special educators identifying parent concerns as a substantial issue (Christmas, 1992).

Implementation

The evaluation of the Utah program to more fully include students labelled as having severe disabilities found that the mean level of model implementation for second-year teachers was 95% across all components of the program (McDonnell et al., 1991). In a program designed to more fully include students labelled as having severe disabilities, the context of systemic reform in state affected data collection and allowed a focus on both integration and inclusion that had been unanticipated in the original evaluation design (Ferguson, 1992).

Overall

Finally, some evaluations make global statements about the success of the programs. In a program to more fully include students described as having severe disabilities in the general classroom, three broad conclusions were reached by the evaluation team. First, integration does not work but inclusion does. Second, integration does not *work* but can be a step on the way to inclusion. And third, inclusion only works well in the context of *reinvented* schools (Ferguson, 1992). In addition, this evaluation found a *strong school effect*: schools that were "learning new stuff" had more powerful effects on processes and outcomes than schools with social integration purposes only. They further concluded that inclusion requires systemic change so that barriers and norms separating regular and special educators break down. This encourages a climate in which they can reinvent learning and schooling, and create environments that foster a sense of belonging for everyone (Ferguson, 1992).

References

Baird, W. (1990). Evaluation of "Blueprint for the future": Literature review and evaluation "tool kit." Scarborough (Ontario) Board of Education. Research Centre Program Department. ED 354 246

Burrello, L. C., & Wright, P. T. (Eds.). (Winter, 1993). Strategies for inclusion of behaviorally challenging students. **The Principal Letters**, 10. ED 358 654

Chase, V., & Pope, E. (1993). Model for mainstreaming: The synergistic approach. Paper presented at the Learning Disabilities of America Conference (San Francisco, CA, February 24-27, 1993). ED 357 594

Christmas, O. L. (1992). The 1992 Michigan non-mandated aide pilot project. Michigan State Department of Education, Lansing, MI. ED 350 745

Co-teaching: Regular education/special education and co-teaching reference guide. (1991). L'Anse Creuse Public Schools, Mt. Clemens, MI. Michigan State Department of Education, Lansing, MI. ED 335 851

Ferguson, D. L. (1992). Regular Class Participation System (RCPS). A Final Report. University of Oregon, Eugene, Oregon. ED 302 308

Karasoff, P., et al. (1992). Systems change: A review of effective practices. *In* P. Campbell et al, Statewide systems change project for the integration of severely disabled students in California: The PEERS Project (Providing Education for Everyone in Regular Schools), Final Report.

Marwell, B. E. (1990). Integration of students with mental retardation. Summary evaluation report. Madison Public Schools, Madison, Wisconsin. ED 355 727

McDonnell, A., McDonnell, J., Hardman, M, & McCune, G. (1991). Educating students with severe disabilities in their neighborhood school: The Utah elementary integration model. *Remedial and special education, 12*:6, 34-45.

Rogan, P., & Davern, L. (1992). Inclusive Education Project: A building-based approach to developing classroom and school models that include students with severe handicaps. Final Report. Syracuse City School District, Syracuse, NY. ED 354 712